

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/566,644  
Source: IFWP  
Date Processed by STIC: 2/14/06

# ***ENTERED***



IFWP

## RAW SEQUENCE LISTING

DATE: 02/14/2006

PATENT APPLICATION: US/10/566,644

TIME: 13:48:14

Input Set : F:\Sequence listing (12810-00197-US).txt

Output Set: N:\CRF4\02142006\J566644.raw

3 <110> APPLICANT: Plesch, Gunnar  
 4 Puzio, Piotr  
 5 Blau, Astrid  
 6 Looser, Ralf  
 7 Wendel, Birgit  
 8 Kamlage, Beate  
 9 Chardonnens, Agnes  
 10 Shirley, Amber  
 11 Wang, Xi-Qing  
 12 Sarria-Millan, Rodrigo  
 13 McKersie, Bryan  
 14 Chen, Ruoying  
 16 <120> TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF FINE CHEMICALS  
 18 <130> FILE REFERENCE: 12810-00197-US  
 C--> 20 <140> CURRENT APPLICATION NUMBER: US/10/566,644  
 C--> 20 <141> CURRENT FILING DATE: 2006-01-31  
 20 <150> PRIOR APPLICATION NUMBER: EP 03016672.2  
 21 <151> PRIOR FILING DATE: 2003-08-01  
 23 <150> PRIOR APPLICATION NUMBER: PCT/US2004/11887  
 24 <151> PRIOR FILING DATE: 2004-04-15  
 26 <160> NUMBER OF SEQ ID NOS: 400  
 28 <170> SOFTWARE: PatentIn version 3.3  
 31 <210> SEQ ID NO: 1  
 32 <211> LENGTH: 579  
 33 <212> TYPE: DNA  
 34 <213> ORGANISM: Saccharomyces cerevisiae  
 36 <220> FEATURE:  
 37 <221> NAME/KEY: CDS  
 38 <222> LOCATION: (1)..(579)  
 40 <400> SEQUENCE: 1  
 41 atg tct gaa aag gcc gtt aga agg aaa ctt gtt att att ggt gat ggt 48  
 42 Met Ser Glu Lys Ala Val Arg Arg Lys Leu Val Ile Ile Gly Asp Gly  
 43 1 5 10 15  
 45 gct tgt ggc aag acc tct tta cta tat gta ttt aca tta gga aaa ttc 96  
 46 Ala Cys Gly Lys Thr Ser Leu Leu Tyr Val Phe Thr Leu Gly Lys Phe  
 47 20 25 30  
 49 cct gaa caa tat cat ccg aca gtg ttc gag aat tat gtc act gat tgc 144  
 50 Pro Glu Gln Tyr His Pro Thr Val Phe Glu Asn Tyr Val Thr Asp Cys  
 51 35 40 45  
 53 aga gtt gac gga ata aaa gtg tcc tta acg ctc tgg gat aca gcg gga 192  
 54 Arg Val Asp Gly Ile Lys Val Ser Leu Thr Leu Trp Asp Thr Ala Gly  
 55 50 55 60  
 57 caa gag gaa tat gaa cgt tta cgt cca ttc tca tat tca aaa gca gat 240

## RAW SEQUENCE LISTING

DATE: 02/14/2006

PATENT APPLICATION: US/10/566,644

TIME: 13:48:14

Input Set : F:\Sequence listing (12810-00197-US).txt

Output Set: N:\CRF4\02142006\J566644.raw

```

58 Gln Glu Glu Tyr Glu Arg Leu Arg Pro Phe Ser Tyr Ser Lys Ala Asp
59 65              70              75              80
61 ata ata tta att ggg ttt gct gta gac aat ttt gaa tca cta att aac      288
62 ile ile leu ile gly phe ala val asp asn phe glu ser leu ile asn
63              85              90              95
65 gca agg acg aaa tgg gcg gat gag gca tta cga tat tgt cct gac gca      336
66 ala arg thr lys trp ala asp glu ala leu arg tyr cys pro asp ala
67              100             105             110
69 cca atc gtt ctt gta ggc ttg aaa aaa gat ttg agg caa gaa gcc cat      384
70 pro ile val leu val gly leu lys lys asp leu arg gln glu ala his
71              115             120             125
73 ttt aaa gag aat gct acc gat gaa atg gtt ccc att gaa gat gca aaa      432
74 phe lys glu asn ala thr asp glu met val pro ile glu asp ala lys
75              130             135             140
77 caa gtt gca agg gcc att ggg gcc aag aaa tac atg gaa tgt agt gca      480
78 gln val ala arg ala ile gly ala lys lys tyr met glu cys ser ala
79 145              150              155              160
81 ctg act ggt gag ggt gtg gat gat gtc ttt gaa gta gct aca aga acc      528
82 leu thr gly glu gly val asp asp val phe glu val ala thr arg thr
83              165              170              175
85 agt ttg ctt atg aag aag gaa cca ggg gct aac tgt tgc ata att tta      576
86 ser leu leu met lys lys glu pro gly ala asn cys cys ile ile leu
87              180             185             190
89 taa      579
92 <210> SEQ ID NO: 2
93 <211> LENGTH: 192
94 <212> TYPE: PRT
95 <213> ORGANISM: Saccharomyces cerevisiae
97 <400> SEQUENCE: 2
99 Met Ser Glu Lys Ala Val Arg Arg Lys Leu Val Ile Ile Gly Asp Gly
100 1              5              10              15
103 Ala Cys Gly Lys Thr Ser Leu Leu Tyr Val Phe Thr Leu Gly Lys Phe
104              20              25              30
107 Pro Glu Gln Tyr His Pro Thr Val Phe Glu Asn Tyr Val Thr Asp Cys
108              35              40              45
111 Arg Val Asp Gly Ile Lys Val Ser Leu Thr Leu Trp Asp Thr Ala Gly
112              50              55              60
115 Gln Glu Glu Tyr Glu Arg Leu Arg Pro Phe Ser Tyr Ser Lys Ala Asp
116 65              70              75              80
119 ile ile leu ile gly phe ala val asp asn phe glu ser leu ile asn
120              85              90              95
123 ala arg thr lys trp ala asp glu ala leu arg tyr cys pro asp ala
124              100             105             110
127 pro ile val leu val gly leu lys lys asp leu arg gln glu ala his
128              115             120             125
131 phe lys glu asn ala thr asp glu met val pro ile glu asp ala lys
132              130             135             140
135 gln val ala arg ala ile gly ala lys lys tyr met glu cys ser ala
136 145              150             155             160

```

## RAW SEQUENCE LISTING

DATE: 02/14/2006

PATENT APPLICATION: US/10/566,644

TIME: 13:48:14

Input Set : F:\Sequence listing (12810-00197-US).txt

Output Set: N:\CRF4\02142006\J566644.raw

```

139 Leu Thr Gly Glu Gly Val Asp Asp Val Phe Glu Val Ala Thr Arg Thr
140                               165                               170                               175
143 Ser Leu Leu Met Lys Lys Glu Pro Gly Ala Asn Cys Cys Ile Ile Leu
144                               180                               185                               190
147 <210> SEQ ID NO: 3
148 <211> LENGTH: 600
149 <212> TYPE: DNA
150 <213> ORGANISM: Oryza sativa
152 <220> FEATURE:
153 <221> NAME/KEY: CDS
154 <222> LOCATION: (1)..(600)
156 <400> SEQUENCE: 3
157 atg ggt tgc tcc tcc tcc gtg cca gct cga agc act gga ggg ttg aac      48
158 Met Gly Cys Ser Ser Val Pro Ala Arg Ser Thr Gly Gly Leu Asn
159 1                               5                               10                               15
161 aat att agc aac gat aac tcc gct act gat tca aag gac ttg cgt gct      96
162 Asn Ile Ser Asn Asp Asn Ser Ala Thr Asp Ser Lys Asp Leu Arg Ala
163                               20                               25                               30
165 aag ttg gta ttg ctt ggt gac tct ggt gta ggg aaa agt tgc att gtt      144
166 Lys Leu Val Leu Leu Gly Asp Ser Gly Val Gly Lys Ser Cys Ile Val
167                               35                               40                               45
169 ctt cgc ttt gtt cgt ggt cag ttt gat ccc act tcc aag gta act gtc      192
170 Leu Arg Phe Val Arg Gly Gln Phe Asp Pro Thr Ser Lys Val Thr Val
171                               50                               55                               60
173 ggt gca tca ttt tta tca caa aca ttg gct ttg gag gac tca aca ata      240
174 Gly Ala Ser Phe Leu Ser Gln Thr Leu Ala Leu Glu Asp Ser Thr Ile
175 65                               70                               75                               80
177 gtg aaa ttt gaa ata tgg gat acc gct gga caa gag agg tat gct gcc      288
178 Val Lys Phe Glu Ile Trp Asp Thr Ala Gly Gln Glu Arg Tyr Ala Ala
179                               85                               90                               95
181 ttg gca cct ctt tac tac aga gga gct gct gct gca gtt gtt gtc tac      336
182 Leu Ala Pro Leu Tyr Tyr Arg Gly Ala Ala Ala Ala Val Val Val Tyr
183                               100                              105                              110
185 gac ata act agt cca gaa tca ttt agc aaa gca caa tac tgg gtg aag      384
186 Asp Ile Thr Ser Pro Glu Ser Phe Ser Lys Ala Gln Tyr Trp Val Lys
187                               115                              120                              125
189 gaa ctt caa aaa cat ggt agt cct gat att atc atg gtt ttg gtt ggt      432
190 Glu Leu Gln Lys His Gly Ser Pro Asp Ile Ile Met Val Leu Val Gly
191                               130                              135                              140
193 aat aaa gct gat cta cat gaa aat cga cat gta tct tct cag gaa gca      480
194 Asn Lys Ala Asp Leu His Glu Asn Arg His Val Ser Ser Gln Glu Ala
195 145                              150                              155                              160
197 caa gag tat gca gag aag aat aat atg gtt ttc atc gag aca tca gca      528
198 Gln Glu Tyr Ala Glu Lys Asn Asn Met Val Phe Ile Glu Thr Ser Ala
199                               165                               170                               175
201 aag aca gct gat aat ata aac caa gta ttt gag gaa att gcg aag agg      576
202 Lys Thr Ala Asp Asn Ile Asn Gln Val Phe Glu Glu Ile Ala Lys Arg
203                               180                               185                               190
205 ttg ccc agg cca acg gcg tct tga      600

```

## RAW SEQUENCE LISTING

DATE: 02/14/2006

PATENT APPLICATION: US/10/566,644

TIME: 13:48:14

Input Set : F:\Sequence listing (12810-00197-US).txt

Output Set: N:\CRF4\02142006\J566644.raw

```

206 Leu Pro Arg Pro Thr Ala Ser
207      195
210 <210> SEQ ID NO: 4
211 <211> LENGTH: 199
212 <212> TYPE: PRT
213 <213> ORGANISM: Oryza sativa
215 <400> SEQUENCE: 4
217 Met Gly Cys Ser Ser Ser Val Pro Ala Arg Ser Thr Gly Gly Leu Asn
218 1      5      10      15
221 Asn Ile Ser Asn Asp Asn Ser Ala Thr Asp Ser Lys Asp Leu Arg Ala
222      20      25      30
225 Lys Leu Val Leu Leu Gly Asp Ser Gly Val Gly Lys Ser Cys Ile Val
226      35      40      45
229 Leu Arg Phe Val Arg Gly Gln Phe Asp Pro Thr Ser Lys Val Thr Val
230      50      55      60
233 Gly Ala Ser Phe Leu Ser Gln Thr Leu Ala Leu Glu Asp Ser Thr Ile
234 65      70      75      80
237 Val Lys Phe Glu Ile Trp Asp Thr Ala Gly Gln Glu Arg Tyr Ala Ala
238      85      90      95
241 Leu Ala Pro Leu Tyr Tyr Arg Gly Ala Ala Ala Ala Val Val Val Tyr
242      100     105     110
245 Asp Ile Thr Ser Pro Glu Ser Phe Ser Lys Ala Gln Tyr Trp Val Lys
246      115     120     125
249 Glu Leu Gln Lys His Gly Ser Pro Asp Ile Ile Met Val Leu Val Gly
250      130     135     140
253 Asn Lys Ala Asp Leu His Glu Asn Arg His Val Ser Ser Gln Glu Ala
254 145     150     155     160
257 Gln Glu Tyr Ala Glu Lys Asn Asn Met Val Phe Ile Glu Thr Ser Ala
258      165     170     175
262 Lys Thr Ala Asp Asn Ile Asn Gln Val Phe Glu Glu Ile Ala Lys Arg
263      180     185     190
266 Leu Pro Arg Pro Thr Ala Ser
267      195
270 <210> SEQ ID NO: 5
271 <211> LENGTH: 648
272 <212> TYPE: DNA
274 <213> ORGANISM: Oryza sativa
276 <220> FEATURE:
277 <221> NAME/KEY: CDS
278 <222> LOCATION: (1)..(648)
280 <400> SEQUENCE: 5
281 atg gcg tcc agc gcg tgc cgg ttc atc aag tgc gtc acg gtc ggg gac      48
282 Met Ala Ser Ser Ala Ser Arg Phe Ile Lys Cys Val Thr Val Gly Asp
283 1      5      10      15
285 ggc gcc gtc ggc aag acc tgc atg ctc atc tgc tac acc agc aac aag      96
286 Gly Ala Val Gly Lys Thr Cys Met Leu Ile Cys Tyr Thr Ser Asn Lys
287      20      25      30
289 ttc ccc act gat tac gta ccc act gtt ttt gac aat ttc agt gca aac      144
290 Phe Pro Thr Asp Tyr Val Pro Thr Val Phe Asp Asn Phe Ser Ala Asn

```

## RAW SEQUENCE LISTING

DATE: 02/14/2006

PATENT APPLICATION: US/10/566,644

TIME: 13:48:14

Input Set : F:\Sequence listing (12810-00197-US).txt

Output Set: N:\CRF4\02142006\J566644.raw

```

291          35          40          45
293 gtg gtg gtc gac ggc acc acg gtg aat ttg ggt ctc tgg gat act gca      192
294 Val Val Val Asp Gly Thr Thr Val Asn Leu Gly Leu Trp Asp Thr Ala
295          50          55          60
297 ggg cag gaa gat tac aac aga ttg agg ccg cta agc tac cgt ggc gcc      240
298 Gly Gln Glu Asp Tyr Asn Arg Leu Arg Pro Leu Ser Tyr Arg Gly Ala
299 65          70          75          80
301 gat gtc ttt gtg ctt gcc ttc tcc cta gtg agc cga gct agc tat gag      288
303 Asp Val Phe Val Leu Ala Phe Ser Leu Val Ser Arg Ala Ser Tyr Glu
304          85          90          95
306 aat gtc atg aag aag tgg tta cca gag ctt cag cat tat gca cca ggg      336
307 Asn Val Met Lys Lys Trp Leu Pro Glu Leu Gln His Tyr Ala Pro Gly
308          100          105          110
310 gtg cca att gtg ttg gtt ggg acc aaa ttg gat ctt cgt gaa gat aaa      384
311 Val Pro Ile Val Leu Val Gly Thr Lys Leu Asp Leu Arg Glu Asp Lys
312          115          120          125
314 cac tac tta ctt gac cat cct agc ttg gtg cct gtg act aca gca cag      432
315 His Tyr Leu Leu Asp His Pro Ser Leu Val Pro Val Thr Thr Ala Gln
316          130          135          140
318 gga gag gaa ctc cgc aag cac att ggc gca acg tgt tac atc gaa tgc      480
319 Gly Glu Glu Leu Arg Lys His Ile Gly Ala Thr Cys Tyr Ile Glu Cys
320 145          150          155          160
322 agc tca aag aca cag cag aat gta aaa gct gtg ttt gat gct gcc atc      528
323 Ser Ser Lys Thr Gln Gln Asn Val Lys Ala Val Phe Asp Ala Ala Ile
324          165          170          175
326 aag gta gta atc aag cct cca aca aag cag agg gac agg aag aag aag      576
327 Lys Val Val Ile Lys Pro Pro Thr Lys Gln Arg Asp Arg Lys Lys Lys
328          180          185          190
330 aaa aca cgg cgc gga tgt tct ttc ttc tgc aag ggt gtc atg tcc aga      624
331 Lys Thr Arg Arg Gly Cys Ser Phe Cys Lys Gly Val Met Ser Arg
332          195          200          205
334 aga agg cta gta tgc ttc aag tga      648
337 Arg Arg Leu Val Cys Phe Lys
338          210          215
341 <210> SEQ ID NO: 6
342 <211> LENGTH: 215
343 <212> TYPE: PRT
344 <213> ORGANISM: Oryza sativa
346 <400> SEQUENCE: 6
348 Met Ala Ser Ser Ala Ser Arg Phe Ile Lys Cys Val Thr Val Gly Asp
349 1          5          10          15
352 Gly Ala Val Gly Lys Thr Cys Met Leu Ile Cys Tyr Thr Ser Asn Lys
353          20          25          30
356 Phe Pro Thr Asp Tyr Val Pro Thr Val Phe Asp Asn Phe Ser Ala Asn
357          35          40          45
360 Val Val Val Asp Gly Thr Thr Val Asn Leu Gly Leu Trp Asp Thr Ala
361          50          55          60
365 Gly Gln Glu Asp Tyr Asn Arg Leu Arg Pro Leu Ser Tyr Arg Gly Ala
366 65          70          75          80

```

**RAW SEQUENCE LISTING ERROR SUMMARY**  
PATENT APPLICATION: US/10/566,644

DATE: 02/14/2006  
TIME: 13:48:15

Input Set : F:\Sequence listing (12810-00197-US).txt  
Output Set: N:\CRF4\02142006\J566644.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:47; Xaa Pos. 15,16,17,18,20,23  
Seq#:48; Xaa Pos. 2,4,5,8,9,10,13,14,15,16,17,18,19,20,21,22,23,25,26,27,28  
Seq#:48; Xaa Pos. 29,30  
Seq#:50; Xaa Pos. 1,9,10,11,12,14,16  
Seq#:52; Xaa Pos. 1,2,4,7  
Seq#:141; N Pos. 980,981  
Seq#:231; N Pos. 27,94  
Seq#:231; Xaa Pos. 2,24  
Seq#:232; Xaa Pos. 2,24  
Seq#:289; N Pos. 957  
Seq#:387; N Pos. 55  
Seq#:397; Xaa Pos. 2,4,5,9,10,14,15,16,17,18,19,20  
Seq#:398; Xaa Pos. 1,3,4,6,9,11,12,13  
Seq#:399; Xaa Pos. 1,2,3,4,5,6,9,12,13,14,15,16,17,19,20,21  
Seq#:400; Xaa Pos. 2,4,10,18,19,20,21,22,23,24,25,26,27,28,29,30,31,34,35  
Seq#:400; Xaa Pos. 37,42,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,60,71  
Seq#:400; Xaa Pos. 79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97  
Seq#:400; Xaa Pos. 98,99,100,103,104,105,106,109,110,111,112,113,115,116  
Seq#:400; Xaa Pos. 118,119,120,123,126,127,128,129,130,131,133,134,136,137  
Seq#:400; Xaa Pos. 141,143,147

## VERIFICATION SUMMARY

DATE: 02/14/2006

PATENT APPLICATION: US/10/566,644

TIME: 13:48:15

Input Set : F:\Sequence listing (12810-00197-US).txt

Output Set: N:\CRF4\02142006\J566644.raw

L:20 M:270 C: Current Application Number differs, Replaced Current Application No  
L:20 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:2963 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:0  
M:341 Repeated in SeqNo=47  
L:2988 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:0  
M:341 Repeated in SeqNo=48  
L:3061 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:0  
L:3114 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:0  
L:7225 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:141 after pos.:929  
L:11581 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:231 after pos.:0  
M:341 Repeated in SeqNo=231  
L:11647 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:232 after pos.:0  
M:341 Repeated in SeqNo=232  
L:14406 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:289 after pos.:912  
L:19025 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:387 after pos.:0  
L:19477 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:397 after pos.:0  
M:341 Repeated in SeqNo=397  
L:19533 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:398 after pos.:0  
L:19626 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:399 after pos.:0  
M:341 Repeated in SeqNo=399  
L:19650 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:400 after pos.:0  
M:341 Repeated in SeqNo=400